

# The B- to A-DNA Transition and the Reorganization of Solvent at the DNA Surface

Nina Pastor

Facultad de Ciencias, Universidad Autónoma del Estado de Morelos, Cuernavaca, Morelos, México

**ABSTRACT** DNA geometry depends on relative humidity. Using the CHARMM22 force field to push B-DNA to A-DNA, a molecular dynamics simulation of a mixed-sequence 24-basepair DNA double-stranded oligomer, starting from B-DNA, was carried out to explore both the mechanism of the transition and the evolution of hydration patterns on the surface of DNA. Over the 11-ns trajectory, the transition recapitulates the *slide-first*, *roll-later* mechanism, is opposed by DNA electrostatics, and is favored by an increasing amount of condensed sodium ions. Hydration was characterized by counting the hydrogen bonds between water and DNA, and by the number of water bridges linking two DNA atoms. The number of hydrogen bonds between water and DNA remains constant during the transition, but there is a 40% increase in the number of water bridges, in agreement with the principle of economy of hydration. Water bridges emerge as delicate sensors of both structure and dynamics of DNA. Both local flexibility and the frustration of the water network on the surface of DNA probably account for the low populations and short residence times of the bridges, and for the lubricant role of water in ligand-DNA interactions.

## INTRODUCTION

The conformation of DNA is sensitive to hydration (Franklin and Gosling, 1953). The fiber diffraction pattern of DNA changes with the relative humidity at which the samples are prepared. Diffraction patterns were interpreted as the canonical A- and B-forms of DNA, the low and high water activity conformations, respectively. The B-form is a slender helix, with sugars in both the *E*- and *S*-conformations, and essentially zero *slide* and *inclination*; the interphosphate distances in this form are 6.5 Å, and the major groove is wider than the minor groove. On the other hand, the A-form is a thicker helix, with sugars practically locked in the *N*-conformation, negative *slide* and *inclination*, interphosphate distances of 5.5 Å, a shallow and wide minor groove, and a deep and narrow major groove. Each helix type presents specific hydration patterns, discussed in detail by Berman and Schneider (1999). B-DNA displays either a hydration spine (when the minor groove is particularly narrow) or a pair of hydration ribbons (with wider minor grooves), and independent hydration of phosphate groups. A-DNA has a clearer hydration structure in the major groove, and successive phosphate groups along each strand of the DNA molecule are bridged by water molecules, lining the major groove. The transition from B- to A-DNA is cooperative, and can be driven by an increase in ethanol, trifluoroethanol (TFE), or salt in the solution. A recent report (Jose and Porschke, 2004) describes the transition driven by dipolar stretching, and suggests the presence of a significant energy barrier separating the two allomorphs. The mecha-

nism of this transition, and the specific role that solvent plays in it, are not clear. The driving force for the transition has been the focus of a large number of studies. Since molecular dynamics simulations offer detailed insight into local and global changes involving both the DNA and the solvent, the present study looked at elements of the mechanisms proposed as key drivers of the B- to A-DNA transition.

The principle of economy of hydration, as stated by Saenger et al. (1986), was put forth to explore possible driving forces for conformational transitions in DNA. Using the DNA crystal structures available at the time, these authors determined that bases and sugars are equally well hydrated in A-, B-, and Z-DNA, with particular ordered patterns of hydration that stabilize each form, as discussed above. In contrast, differences in sugar-ring pucker lead to different intrastrand *P-P* distances and to the possibility of forming water bridges between O1P-O1P and O1P-O2P (*i*, *i*-1) in A-DNA but not in B-DNA. This leads to fewer water molecules required to hydrate A-DNA than B-DNA, and therefore to economy of hydration.

It was also recognized early on that not all DNA sequences are amenable to the B- to A-transition. Statistical analyses linking alcohol-induced B- to A-transitions to differences in solvent-free energies (Basham et al., 1995) or to sequence frequencies (Tolstorukov et al., 2001), have come up with different dimer and trimer scales for the A-philicity or B-philicity of particular DNA sequences. Although there is no universal scale, most scales agree on making G-rich sequences A-philic, and A-rich sequences B-philic. A host of rationales have been given to explain the trends in the scales. For example, the preference for the B-form of both AA/TT and AAN/NTT (*N* = A, T, C, G) is explained by the hydration spine in the minor groove, by thymine methyl-sugar interactions in the major groove, and by bifurcated

Submitted December 20, 2004, and accepted for publication March 1, 2005.

Address reprint requests to Nina Pastor, Facultad de Ciencias, Universidad Autónoma del Estado de Morelos, Av. Universidad 1001, Col. Chamilpa, 62210 Cuernavaca, Morelos, México. Tel.: 52-777-3-29-70-20; Fax: 52-777-3-29-70-40; E-mail: nina@servm.fc.uaem.mx.

© 2005 by the Biophysical Society

0006-3495/05/05/3262/14 \$2.00

doi: 10.1529/biophysj.104.058339

H-bonds in the major groove (Tolstorukov et al., 2001). The preference for the A-form of GC basepairs is ascribed to the ability of cytosine to switch to C3'endo, in agreement with recent QM calculations (Foloppe and MacKerell, 1999). In all cases, hydration is assumed to play a major role in either facilitating or hindering the transition.

Computational exploration of the free energy change associated with the transition has been done using the free-energy component analysis. This kind of analysis does not require knowledge of the pathway connecting the two states. Srinivasan et al. (1998) found that salt effects favor A-DNA by a modest amount, whereas phosphate-phosphate repulsion favors B-DNA; solvation counteracts this advantage, though. Added salt further favors A-DNA, whereas stacking favors B-DNA. Regarding sequence preferences, they found that G-rich sequences have a smaller free energy difference than A-rich sequences in favor of B-DNA, in agreement with the trends observed experimentally. Jayaram et al. (1998) carried out MD simulations of the *EcoRI* dodecamer in water and an ethanol/water mixture. They found that the explicit organization of mobile counterions around DNA is key to the transition. The preference for B- over A-DNA in water is due in this analysis to a combination of reduced phosphate repulsion and improved hydration. In ethanol, A- wins over B-DNA because of reduced electrostatics (the water/ethanol mixture has a lower dielectric constant than water, which results in a lower solvation energy difference between the A- and B-forms of DNA) and higher counterion condensation. In both cases, it is the salt that apparently pays for the internal energy cost of bringing the phosphate groups closer together in A-DNA compared to B-DNA. Other calculations with more recent versions of the AMBER force field have arrived at the same conclusions (Cheatham et al., 1998). It must be noted that the AMBER force field (Cornell et al., 1995) does not produce stable A-DNA conformations in pure water and high relative humidity (see below), so all the simulations starting from the canonical A-DNA conformation used restraints to keep it that way.

In an interesting computational experiment, Mazur (2003) described reversible transitions between A- and B-DNA of a single double-helix in a water drop with sodium, by reducing drop size. Using the AMBER95 force field (Cornell et al., 1995) in internal coordinate molecular dynamics, and run lengths from 5 to 25 ns, the simulations started from canonical B-DNA, and upon reduction of the drop size, underwent the transition to A-DNA. The structure found at the lowest hydration level was then rehydrated and found to return to B-DNA. The proposed mechanism for the B- to A-DNA transition is that direct electrostatic interactions between mobile metal ions and phosphate groups across the major groove lead to DNA bending, then to sugar repuckering, and therefore to the transition.

In recent years, intermediates between A- and B-DNA forms have been reported experimentally, both from x-ray crystallography and NMR. Furthermore, it has been pro-

posed that some DNA binding proteins recognize an intermediate structure between A- and B-DNA (Jones et al., 1999; Lu et al., 2000; Neklodova and Pabo, 1994; Olson et al., 1998). For example, the structure of the oligomer GCATATGATAG (Tonelli et al., 1998), which happens to be a bacterial promoter, was shown, in an NMR study, to be such an intermediate. Crystallographic studies reported an intermediate conformation for a GC-rich oligomer, CATGGGCCCATG, in which the G-tract achieves the low *slide* characteristic of A-DNA, but not positive *roll*, and is therefore blocked halfway in the transition (Ng et al., 2000). In this work, it is proposed that electrostatic interactions between consecutive Gs in B-DNA create a destabilizing tension that increases cooperatively and drives the transition. In a subsequent study, the same group (Ng and Dickerson, 2002) performed a careful analysis of the hydration patterns present in the intermediate, and suggested that the different hydration of GC and AT basepairs provide a physical basis for solvent-dependent facilitation of the B-to-A transition by GC basepairs. The latter are more heavily hydrated than the flanking AT basepairs. Vargason et al. (2001) reported 13 intermediates of the hexamer GGCGCC with methylations and brominations. These intermediates were classified according to the backbone dihedrals  $\delta$  and  $\chi$ , and they propose that changes in *slide* precede those in *inclination*, in agreement with the *slide-first, roll-later* mechanism proposed (Ng et al., 2000). In contrast to the x-ray-derived data, extensive spectroscopic analysis (Dornberger et al., 2001) of the Ng et al. (2000) dodecamer by FTIR, NMR, and CD, revealed no significant population of *N*-type sugars in the absence of TFE, and a cooperative B- to A-DNA transition in TFE, as opposed to the mixed structure found crystallographically. The experimental data were backed by various MD simulations of the dodecamer, carried out with three different versions of AMBER potentials (*parm95*, *parm98*, and *parm99*) starting from canonical A- and B-DNA structures and from the crystal structure, which indicated in all cases a typical B-DNA structure in solution for this dodecamer. This study raises the possibility that the intermediates seen in crystal environments may not occur in solution.

To explore the B- to A-DNA transition in solution, advantage was taken of the known and well-documented tendency of the CHARMM22 (MacKerell et al., 1995) potential to convert DNA oligomers into an A-DNA conformation (Feig and Pettitt, 1997, 1998; Yang and Pettitt, 1996), regardless of the sequence of the oligomer. A mixed sequence with all possible basepair steps was chosen for the double-stranded DNA 24-mer subject of this study: d(CTGTATCTATATAAAACGGGCTGG) • d(CCAGCCCGTTTTATATAGATACAG), complementing previous studies of the hydration of dinucleotide-repeating DNA (Auffinger and Westhof, 2000, 2001) and block copolymers (Feig and Pettitt, 1997, 1998). This sequence contains a G-tract, which favors the transition, an alternating TA-tract,

which complies with the transition, and an A-tract, which should resist the transition. The purpose of this simulation is to characterize both the transition and the evolution of the hydration patterns on the surface of DNA. Since the transition occurs in high relative humidity, it is clearly driven by the force field and could therefore follow a pathway that is not sensible in terms of the mechanics of nucleic acids reported to date. Following a reasonable pathway during the transition would suggest that the evolution of the hydration patterns on the surface of DNA could also be deemed realistic; otherwise, only the analysis of hydration patterns at the endpoints of the transition would be meaningful. Therefore, the first part of this study is concerned with the mechanics of the transition, and having validated it, the second part relates to the changes in hydration that occur as a function of time. The mixed sequence of the oligomer does not allow for averaging over all basepairs to study hydration patterns. Moreover, as the interest is in the evolution of hydration along the transition, this precludes the use of pseudo-electron density maps to characterize solvation patterns. As the solute is constantly changing its conformation with time, hydration was studied by counting the number of hydrogen bonds to water for each donor and acceptor atom in DNA, and also by counting the number of water bridges simultaneously connecting two DNA atoms. This kind of analysis has been previously used by Bonvin et al. (1998).

## METHODS

### System setup and simulation

The 24-mer, double-stranded DNA helix in canonical B-DNA form was built in QUANTA, with the sequence d(CTGTATCTATATAAAACGGGCTGG) • d(CCAGCCGTTTATATAGATACAG). The 5' terminal phosphate groups were deleted, and 5' and 3' ends were capped with OH groups. Forty-six sodium atoms were added along the O-P-O bisector, at 5 Å of the phosphorus, to make the system electroneutral. The oligonucleotide and sodium ions were solvated in a hexagonal prism box, 112 Å in length and 27.72 Å inside, with 7149 TIP3P water molecules (Jorgensen et al., 1983) in InsightII (Accelrys, San Diego, CA); this amounts to a 14 Å water shell surrounding the DNA duplex, and results in DNA images being at least 24 Å apart through the shortest path between images (the flat sides of the hexagon). Periodic boundary conditions, with a cutoff of 13 Å for image generation, were applied throughout the simulation, which was carried out in the NVE ensemble at 300 K. The Verlet algorithm was used to integrate the equations of motion, using a time step of 1.5 fs for the production phase. Hydrogen atom-heavy atom bond lengths were constrained with SHAKE (Ryckaert et al., 1977), and a dielectric constant of 1 was used. The nonbonded list was updated every 10 steps with a 13 Å atom-based cutoff, and coordinates were saved for further analysis every 50 steps (0.075 ps). The simulation was performed with the CHARMM program (Brooks et al., 1983), with the all-atom version 22 parameters (MacKerell et al., 1995), which properly balance solute-solute, solute-solvent, and solvent-solvent nonbonded interactions (MacKerell et al., 2001). The van der Waals interactions were treated with a shifting function in the range of 10–12 Å, and electrostatics were shifted to a final cutoff at 12 Å, as suggested in the parameter file for CHARMM22. This scheme for treating long-range interactions has been shown to produce equivalent results to the use of particle-mesh Ewald (Norberg and Nilsson, 2000).

The water shell was energy-minimized with four cycles of 500 steps of the steepest-descent (SD) algorithm, followed by four cycles of 500 steps of the adopted-basis Newton-Raphson (ABNR) algorithm, using the default scaling for step size. Water and sodium ions were further minimized with four cycles of 500 SD steps and two cycles of 1000 ABNR steps. At this point, periodic boundary conditions were switched on, and water and sodium ions were minimized with 250 SD steps and 250 ABNR steps. Keeping the DNA oligomer fixed, water and sodium ions were heated to 600 K during 10 ps, and equilibrated during 150 ps with a time step of 2 fs. Water and sodium ions were minimized with 250 SD steps and 250 ABNR steps, followed by the same treatment for the whole system. All atoms were heated to 300 K in 10 ps, and equilibrated during 30 ps with a 2-fs time step and a temperature tolerance of  $\pm 5$  K. The time step was then reduced to 1.5 fs, NVE conditions were applied, and the simulation was propagated for 11.2 ns. It must be noted that shifting to NPT conditions resulted in only modest compression of the cell size (data not shown), suggesting that the simulation does not include inordinate pressures.

As standards of hydration at the endpoints of the transition, two additional simulations were performed with fixed canonical A- and B-DNA conformations of the same oligomer. The fixed B-DNA simulation was prepared in the same way as described above, except that no minimization of the whole system was done. After solvent equilibration at 600 K, and solvent minimization, lattice ABNR minimization was carried out for 500 steps, followed by 250 steps of SD and by another round of 250 steps of ABNR. The NPT ensemble was used for the MD simulation, at 300 K and 1 atm. Pressure control was achieved with the Langevin Piston method (Feller et al., 1995), with PMASS 600.0 and PGAMMA 10.0. Temperature control was enforced with the Nose-Hoover method, with TMASS 1000.0. One-hundred-and-fifty picoseconds were used for equilibration, and this was followed by 1 ns of production. For the fixed A-DNA simulation, the oligomer was built in QUANTA, and sodium was added along the O-P-O bisector at 5 Å from the phosphorus. This generated a string of sodium ions along the major groove. This complex was immersed in a box of equilibrated TIP3 water molecules, to generate a 14 Å water shell around the oligomer. Hexagonal prism periodic boundary conditions were used. At this point, the same protocol for the fixed B-DNA simulation was enforced: lattice ABNR minimization was carried out for 500 steps, followed by 250 steps of SD and by another round of 250 steps of ABNR. The NPT ensemble was used for the MD simulation, with the same parameters used for the fixed B-DNA simulation. One-hundred-and-fifty picoseconds were used for equilibration, and this was followed by 3 ns of production. The first nanosecond was necessary to randomize the position of sodium atoms (possibly equivalent to the 150-ps dynamics at 600 K for the B-DNA oligomer). A comparison of hydration patterns during the second and third nanoseconds showed no further change, compared to the first nanosecond. Only the data from the third nanosecond were used for analysis.

A control simulation in which sodium charges were modified to 0.0 was carried out in the NPT ensemble, starting from the prepared B-DNA system immediately after ion and solvent equilibration at 600 K and minimization. The same protocol was used as for the fixed B-DNA simulation, except that all atoms were allowed to move. This simulation was extended to 4 ns.

### Analysis

Coordinate sets saved every 75 fs were analyzed, excluding the terminal basepairs to reduce end effects. The time resolution is comparable to recent ultrafast spectroscopy experiments on DNA hydration (Pal et al., 2003a,b). As all atoms were allowed to move freely during the simulation, DNA both translated and rotated, hitting the walls of the simulation cell by the fifth nanosecond of simulation. To remove these motions, the trajectory was recentered and reoriented with respect to the starting configuration before analysis.

Structural analysis was done with the *Curves* 3.0 version included in the MD Toolchest 1.0 version (Ravishanker et al., 1989) (backbone dihedral angles; basepair step parameters at the dinucleotide level, local frame of

reference; and basepair parameters for single basepairs), and with CHARMM28 (interatomic distances, root mean-square distance from canonical A- and B-structures, solvent-accessible surface area, energy components). Solvent-accessible surface areas were calculated for the complete DNA oligomer with the analytical Lee and Richards method (Lee and Richards, 1971) included in CHARMM28, using a 1.4 Å radius for the probe, but only the data for the internal 22 basepairs was processed. DNA atoms were assigned to major (C2', H2', H2'', C3', H3', O5', P, O1P, C5, C6, N7, C8, and H8 for all; N6, H61, and H62 for adenine; C4, O4, C5M, H51, H52, H53, and H6 for thymine; O6 for guanine; N4, H41, H42, C4, H5, and H6 for cytosine) and minor grooves (C1', H1', C2', H2', H2'', O3', H4', O4', C5', H5', H5'', O2P, and C2 for all; H2, N3, and C4 for adenine; O2 for thymine; N2, H21, H22, N3, and C4 for guanine; O2 for cytosine). The assignment was verified by visual inspection in RasMol (<http://www.umass.edu/microbio/rasmol/>); C2', H2', and H2'' were assigned to both grooves as their position depends on DNA conformation, pointing to the minor groove in A-DNA and to the major groove in B-DNA. Energy calculations were performed with the 13 Å cutoff for nonbonded list generation and 12 Å cutoff for actual computation of nonbonded interactions, unless otherwise noted.

The distribution of sodium around DNA was calculated within the scheme of proximity analysis (Mehrotra and Beveridge, 1980; Mehrotra et al., 1981), as implemented in MMC (<http://fulcrum.physbio.mssm.edu/~mezei/mmc/mmc.html>). The simulation was broken into 2-ns intervals to improve statistics; therefore, the last nanosecond was not analyzed. For each interval, the calculation was done twice: one considering the whole oligomer as one functional group, and another in which DNA atoms were divided into groups corresponding to the phosphate moiety (P, O1P, O2P, O5', and O3'), the minor groove and the major groove, defined as above but removing the phosphate atoms, for each nucleotide. For each snapshot analyzed (five sets of 27,200 snapshots), 10,000 points were used to estimate the volumes of the proximity area of each functional group. Proximity areas were defined by the bisector along the bond connecting atoms.

Hydration properties were calculated using CHARMM28. An H-bond was considered to exist if the donor-acceptor distance lay within 2.4 Å, with no angle constraints. This allows for detection of bifurcated H-bonds, which normally do not have optimal angle parameters. A water-bridge required

simultaneous formation of two H-bonds by the same water molecule to different DNA atoms. Water bridges between DNA and sodium atoms were also calculated. To estimate the average time a water molecule spent making either an H-bond or a bridge, the total time a particular bridge (or H-bond) existed was divided by the number of water molecules that participated in making that bridge (or H-bond). This does not correspond strictly to the residence time, as no consideration is given to the distribution of individual lifetimes.

## RESULTS AND DISCUSSION

### The structural transition

As expected, a B- to A-DNA transition occurs in the oligomer. The transition is monitored in Fig. 1 by calculating the root mean-square (RMS) difference between each snapshot from the trajectory and either the starting canonical B-DNA conformation, or the canonical A-DNA conformation. As the oligomer is 24-basepairs long, it was broken into three regions to make the comparison more meaningful: *Region 1* contains the six basepairs at the 5' end of the oligomer; *Region 2* contains the following nine basepairs; and *Region 3* is formed by the last seven basepairs (note that the terminal basepairs were not taken into account). As shown in Fig. 1, the three regions undergo the transition, at different rates and to differing extents. As expected from the published A-philicity scales, the G-tract (included in Region 3) is the fastest and most thoroughly converted region (Fig. 1 C). Region 2 (Fig. 1 B) has apparently not completed the transition even after 11 ns (this region contains an A-tract, which is reputed to resist the transition the most), and Region 1 (Fig. 1 A) appears to display an oscillatory behavior. The time required for the transition is physically meaningless, as

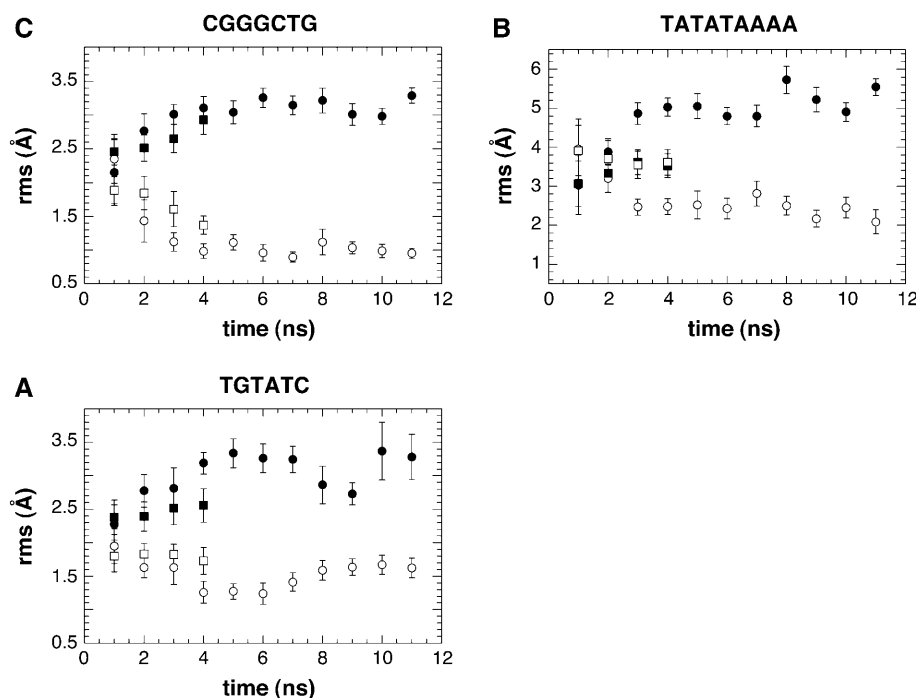


FIGURE 1 Root mean-square deviation of the oligomer structure to canonical A- (*open symbols*) and B-DNA (*solid symbols*), as a function of time. One-nanosecond averages, mean  $\pm$  SD are shown as a function of time for the three regions defined in the text. (A) Region 1; (B) Region 2; and (C) Region 3. (*Circles*, native simulation; *squares*, control simulation with uncharged sodium ions.)

the transition was driven by a persistent bias from the force field. Nevertheless, the fact that different regions of the oligomer comply with this bias at various times suggests that sequence-dependent features are not lost in the simulation.

To explore the driving force of the transition, an internal energy analysis was conducted. Internal energy contains both the bonded and nonbonded interactions in the DNA molecule. As shown in Fig. 2 *A*, the internal energy of DNA becomes less attractive as the transition progresses; each component of the bonded and nonbonded interactions was examined to find the relevant contributions to this trend, and only the electrostatic interactions correlated with the total internal energy behavior, as seen in Fig. 2 *B*. This is entirely expected, as the B- to A-DNA transition entails bringing phosphate groups closer, both within each strand of DNA and also across the major groove. Neither total internal energy nor the electrostatic contribution seems to have reached equilibrium at the time the simulation was stopped. It should be noted that both the bonded and the van der

Waals interactions remained essentially unchanged as the transition occurred. Since DNA opposes the transition, some other part of the system must compensate the energetic cost, as the transition occurs in an NVE simulation. DNA-water interactions were calculated within a 5 Å cutoff of DNA atoms, but this quantity showed no change over time; the reason for this will be discussed in the second section of this report. On the other hand, the DNA-sodium interaction became progressively more attractive along the transition (Fig. 2 *C*), in agreement with free-energy calculations (Jayaram et al., 1998; Srinivasan et al., 1998), and the role proposed by Mazur (2003) for sodium. This interaction is sufficiently attractive to compensate and exceed the unfavorable DNA electrostatics. As would be expected from Manning's counterion-condensation theory (Manning, 1978), the number of condensed sodium ions increases with time due to the decrease in interphosphate distances. This is shown in Fig. 2 *D*. There is an increase in the number of sodium ions located in the proximity regions of both

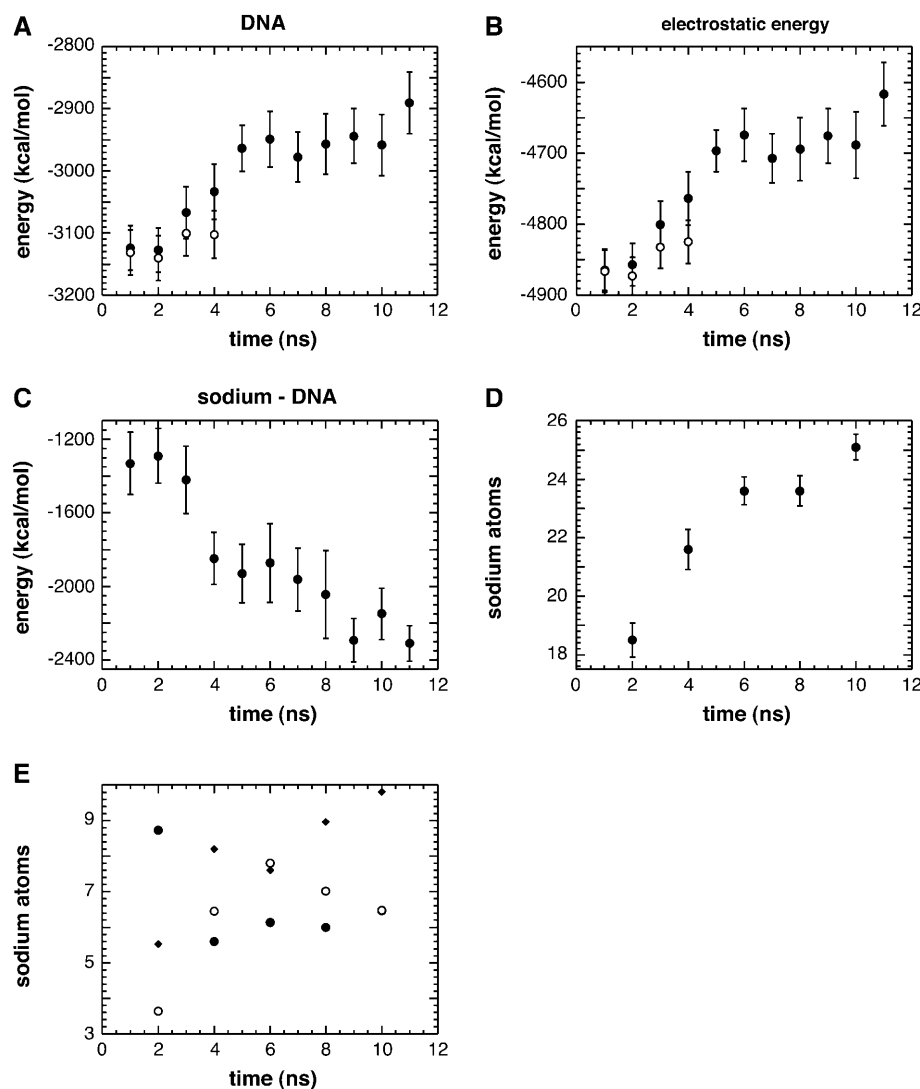


FIGURE 2 Internal energy decomposition for the B- to A-DNA transition. One-nanosecond averages, mean  $\pm$  SD are shown as a function of time. (A) Total internal energy (kcal/mol); (B) electrostatic energy (kcal/mol); and (C) sodium-DNA energy (kcal/mol). (Solid symbols, native simulation; open symbols, control simulation with uncharged sodium ions.) (D) Number of sodium ions within a 6 Å radius of the surface of DNA  $\pm$  twice the mean SD over 2-ns intervals. (E) Number of sodium ions within a 6 Å radius of the surface of DNA over 2-ns intervals, for the minor groove (solid circles), major groove (open circles), and phosphate functional groups (solid diamonds).

phosphates and the major groove, whereas the population of sodium in the minor groove shows first a decrease, followed by a modest increase, as seen in Fig. 2 *E*. Except for a few, long-lived direct contact sodium-DNA interactions at carbonyl oxygens in both grooves, most of the sodium ions interact with DNA through solvent. Approximately 80% of the contacts were made to the phosphate and major groove atoms, in agreement with previous observations (Mazur, 2003).

To assess the importance of the contribution of sodium to the feasibility of the transition, a simulation in which sodium charges were reduced to zero was done. The B- to A-DNA transition seems to be interrupted in the simulation without charged sodium, as evidenced in Figs. 1 and 2. Only Region 3 seems to be undergoing the transition to A-DNA, as seen by the time evolution of the RMS distance to A- and B-DNA (compare Fig. 1, A and C), and both the total and the electrostatic energies remain at more negative values than those found for the original simulation (Fig. 2, A and B). Comparison of geometrical properties from both simulations, shown below, suggests that cations are required to shorten the helix and to compress the major groove, in agreement with previous work with other force fields and simulation protocols (Cheatham and Kollman, 1997a; Mazur, 2003).

To test the hypothesis that guanine-guanine stacking is unfavorable in B-DNA and that the relaxation of this strain drives the transition (Ng et al., 2000), the interaction energy was calculated for both GG/CC basepair steps present in the oligomer (excluding the backbone), as a function of time. The electrostatic component of the interaction energy is constant with time, averaging  $\sim 5$  kcal/mol for both steps; the van der Waals energy is always attractive, averaging  $\sim -17$  kcal/mol, and is also constant with time (data not shown). It appears therefore that with this version of the CHARMM force field, the transition is not driven by an improvement of stacking energies between GC basepairs.

The analysis does not explain what starts the transition, but it suggests that the force-field-induced transition has reasonable energetics, and that a more in-depth analysis of both geometric and hydration properties may provide valuable insight into the nature of the transition.

### On the geometry of DNA along the transition

Crystallographic analysis of intermediates between A- and B-DNA conformations suggested a *slide-first, roll-later* mechanism for the transition (Ng et al., 2000; Vargason et al., 2001). Fig. 3 *A* shows the time evolution of the average *roll* and *slide* for all basepair steps, in 1-ns averages, throughout the transition. *Slide* achieves strongly negative values by 4 ns, whereas *roll* still evolves toward increasingly positive values even after 11 ns. This behavior is consistent with the mechanism proposed by the Dickerson and Ho laboratories, in that *slide* stabilizes first, and *roll* responds later. The fact that *roll* has not stabilized at the end of the simulation

might be related to the upward drift in total energy of DNA (see Fig. 2 *A*), though this remains to be explored in detail. The simulation with zero-charged sodium achieves negative *slide*, which is not as strong as the native simulation, but does not *roll*. It also seems like the transition stopped at the third nanosecond. This situation is reminiscent of the structure of the Ng et al. (2000) oligomer, which presents negative *slide* but little *roll*.

Vargason et al. (2001) classified their 13 intermediates in terms of the correlation between two backbone parameters:  $\delta$  and  $\chi$ . This correlation is described in Saenger's *Principles of Nucleic Acid Structure* as a consequence of the chemical linkage between the sugar and the base, and is also described (Lu et al., 2000) as a discriminating condition between A- and B-type helices. The time evolution of these two parameters is shown in Fig. 3 *B*. For reference,  $\delta$  for B-DNA is  $156^\circ$ , and  $\chi$  is  $264^\circ$ ; the values for A-DNA are  $83^\circ$  and  $207^\circ$ , respectively. It is clear that even within the first nanosecond of simulation, canonical B-DNA values are no longer sampled in the simulation, and as time passes, the system moves further away from B-DNA and converges after 4 ns in the region expected for A-DNA. The correlation between  $\delta$  and  $\chi$  is nearly perfect (0.994). The simulation with uncharged sodium also displays the correlation between  $\delta$  and  $\chi$ , but again seems to have interrupted the transition to A-DNA.

A better descriptor of sugar pucker is the pseudo-rotation angle,  $P$ . Surveys of crystal data have indicated that B-type helices sample predominantly the *E*- and *S*-conformations, but may even extend occasionally to the *N*-region of the pseudo-rotation cycle. For this reason, and because intermediate structures have been found with B-DNA puckers and A-DNA stacking of bases (Trantirek et al., 2000), sugar pucker is no longer considered an unequivocal indicator of helix type. On the other hand, A-type helices appear to have all sugar moieties locked in the *N*-conformation. Sugar puckering correlates with *slide*, in that high negative *slide* requires the *N*-conformation (Dickerson and Ng, 2001). Fig. 3 *C* shows the time evolution of the population in the *N*-conformation ( $P$ -values contained in the interval  $[324^\circ, 36^\circ]$ ), representative of A-DNA, and in the *E* + *S* conformations ( $P$ -values in the interval  $[36^\circ, 216^\circ]$ ), corresponding to B-DNA. It is clear that the *N* population increases at the expense of the *E* + *S* conformations, and that the transition is complete after 4 ns. These data were collected for purines and pyrimidines separately, and purines complete the transition earlier and more thoroughly than pyrimidines (see Supplementary Material). In this respect, CHARMM22 fails to reproduce the reported tendency of cytosines to acquire the *N*-conformation more easily than all other nucleotides in gas phase quantum mechanical calculations, a behavior ascribed to the parameterization of sugar dihedral energetics (Foloppe and MacKerell, 1999). The global behavior displayed by the simulation is consistent with known DNA mechanics. The control simulation with

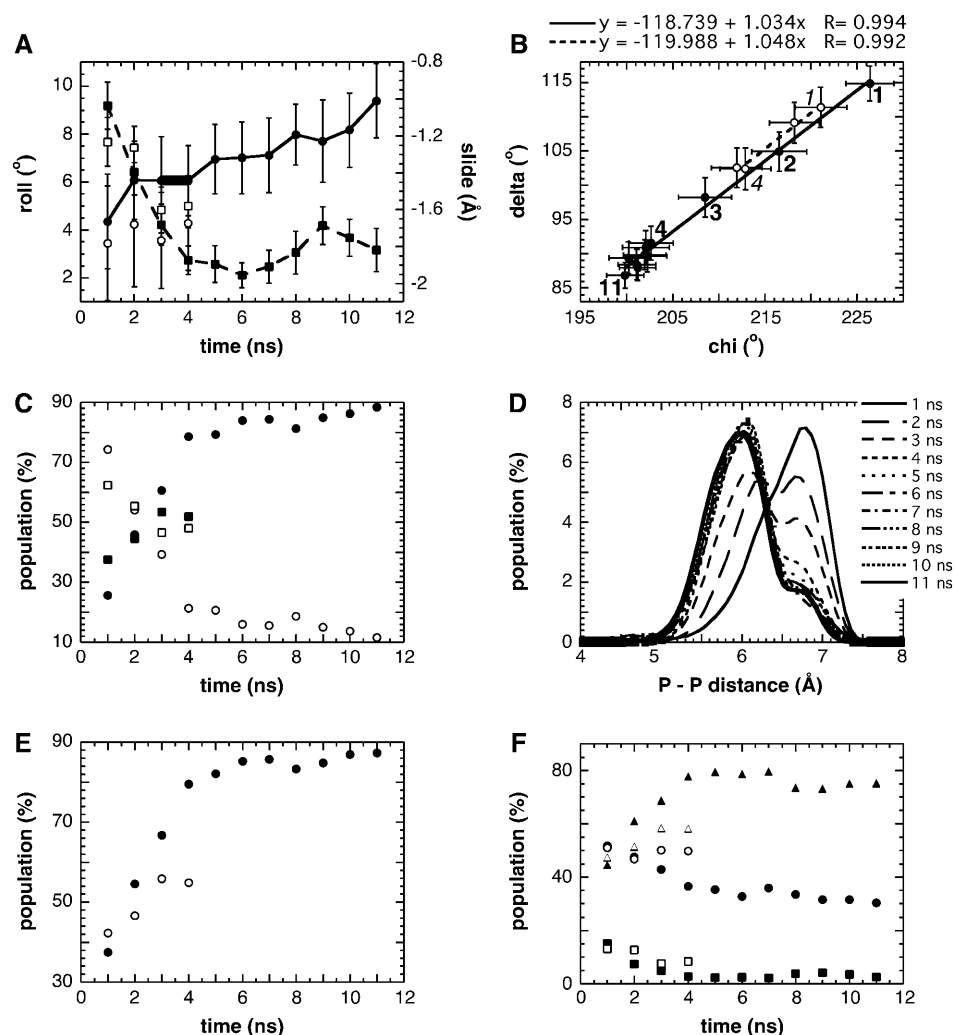


FIGURE 3 Geometrical characterization of the B- to A-DNA transition. One-nanosecond averages, mean  $\pm$  SD are shown as a function of time. (A) Roll (continuous line) and slide (dashed line) versus time. (Solid symbols, native simulation; open symbols, control simulation.) (B) Correlation between  $\delta$  and  $\chi$ ; the corresponding time interval is labeled for 1, 2, 3, 4, and 11 ns (solid symbols and continuous line, native simulation; open symbols and dashed line, control simulation). (C) Population percentage for  $N$  (solid symbols) and  $E + S$  (open symbols) sugar pucker conformations, as a function of time, for 1-ns intervals (circles, native simulation; squares, control simulation). (D) Consecutive  $P-P$  distance histograms for the native simulation, as a function of time, for 1-ns intervals (the first and last nanoseconds are shown in thicker lines). (E) Population percentage for consecutive  $P-P$  distances within 6.4 Å (solid symbols, native simulation; open symbols, control simulation). (F) The values  $\epsilon$ - and  $\zeta$ -conformations as a function of time, for 1-ns intervals (BI in circles, BII in squares, and A in triangles; solid symbols, native simulation; open symbols, control simulation).

uncharged sodium seems to have converged at a 50:50 ratio of sugar pucker in the  $N$ - and the  $E$ - and  $S$ -conformations (Fig. 3 C). Detailed analysis of the behavior of purines and pyrimidines (see Supplementary Material) revealed that 80% of the purines have converted to the  $N$ -conformation, whereas only 20% of the pyrimidines have done so. These pyrimidines are located in Region 3, consistent with the low RMS distance to A-DNA for this region (Fig. 1 C). This low percentage of conversion to the  $N$ -conformation explains the intermediate values of *slide* and  $\chi$  found for the control simulation.

The relative position of consecutive phosphate groups also changes during the B- to A-DNA transition, as a consequence of sugar repuckering. Fig. 3 D shows the time evolution of the distribution of distances between consecutive  $P$  atoms along each strand of the duplex, for 1-ns intervals. At the beginning of the simulation, most of the distances are typical of B-DNA ( $\sim 7$  Å), whereas by 4 ns there is a prominent peak at 6 Å, representative of A-DNA. The population at 7 Å takes 11 ns to disappear; the nucleotides responsible for this

population belong to Regions 1 and 2, defined above. The population of interphosphate distances below 6.4 Å was calculated from these distributions; the cutoff was chosen as the midpoint of the distances at which peaks appeared for the initial and final nanoseconds of the simulation (6.8 and 6 Å, respectively). The time evolution of this population is shown in Fig. 3 E, for both the native and the control simulations. The transition seems to be complete by 4 ns for the native run, and it again seems to be interrupted for the control simulation with zero-charged sodium.

Sugar repuckering also affects the backbone dihedral angles  $\epsilon$  and  $\zeta$ . Fig. 3 F depicts the time-evolution of the populations of BI-, BII-, and A-type conformations, defined as follows: BI  $\epsilon$  [120°–210°],  $\zeta$  [210°–295°]; BII  $\epsilon$  [210°–300°],  $\zeta$  [150°–210°]; A  $\epsilon$  [170°–230°], and  $\zeta$  [270°–330°]. In agreement with previous reports, the BI-conformation is the most common at the start of the simulation (Winger et al., 2000, 1998), but the A-DNA conformation dominates after 4 ns. For the control simulation, the BII-conformation persists almost unaltered after 4 ns, with a modest increase in the

A-DNA conformation, probably at the expense of the BI-conformer.

A commonly used property to classify DNA helices is groove width. In B-DNA, the major groove width is  $\sim 17$  Å, measured between one phosphorus and its partner in the complementary strand, three basepairs downstream; the minor groove width is  $\sim 11$  Å, to a partner that is four basepairs upstream. In A-DNA, the major groove width is  $\sim 7$  Å, six basepairs downstream, whereas the minor groove width is  $\sim 16$  Å, three basepairs upstream. The minimum distance to the phosphorus atoms in the complementary strand was calculated, as a function of time, for each phosphorus atom along one strand of the duplex. This distance represents the narrower groove of the DNA oligomer at that particular site. Phosphorus atoms corresponding to residues 6–19 were analyzed so that both minor and major groove partners were well defined for both structural families. In most cases, the most common minimum distance sampled during the complete simulation corresponded to major groove partners in the A-DNA family, concordant with the fact that the transition from B- to A-DNA took place; the only two exceptions occur in Region 2, for which the most common minimum distance corresponds to the minor groove partners. For the canonical pairs of phosphorus atoms lining the major groove of A-DNA, the first time that the distance between these atoms became 10.5 Å or less during the simulation was recorded. For the phosphates in Region 1 (defined above), major groove compaction occurs at  $\sim 4$  ns; for Region 2 this happens between 5 and 9 ns (the latest being the alternating TA tract), and Region 3 with the G-tract is the first to compact, doing so at 2 ns. This suggests that DNA unwinds first, placing the atoms in the canonical A-DNA register, and then collapses. In this computational experiment, sugars repucker first and then the major groove

closes. Water bridges between sodium and DNA were calculated to look specifically for hydrated sodium ions linking phosphate groups across the major groove. These bridges occur during 1.5% of the simulation time, and thus are unlikely to represent a major stabilizing interaction for the compacting major groove; the effect of sodium would appear to be mediated by the increase in sodium concentration both near the phosphate groups and in the major groove (Fig. 2 E), rather than by specific interactions at particular sites. In the control simulation, the distance of closest approach between phosphates always corresponded to the minor groove, and most contacts never came within 10.5 Å. This, together with the small population of interphosphate distances within 6.4 Å (Fig. 3 E), explains the more attractive total and electrostatic energies found for the control simulation compared to the native one (Fig. 2, A and B).

A consequence of all these conformational changes is the modification of the solvent-accessible surface area of the DNA oligomer, depicted in Fig. 4 A. There is a decrease of  $\sim 300$  Å<sup>2</sup> along the transition, corresponding to  $\sim 30$  water molecules (assuming 10 Å<sup>2</sup> for each water molecule; Parsegian et al., 1995). This agrees with the principle of economy of hydration, in the sense that A-DNA requires fewer water molecules to cover its surface than B-DNA. It must be noted that the translation between accessible surface area and water molecules is an approximation, more so for rugged and curved surfaces like the one on DNA, so these numbers just indicate a reasonable trend. Following known features of A- and B-DNA, the minor groove area increases (Fig. 4 B), whereas the major groove area decreases (Fig. 4 C) with time; this trend agrees with the mechanics of DNA recognition by  $\alpha$ -helices (Lu et al., 2000), resulting in intermediate structures between A- and B-DNA. It has been

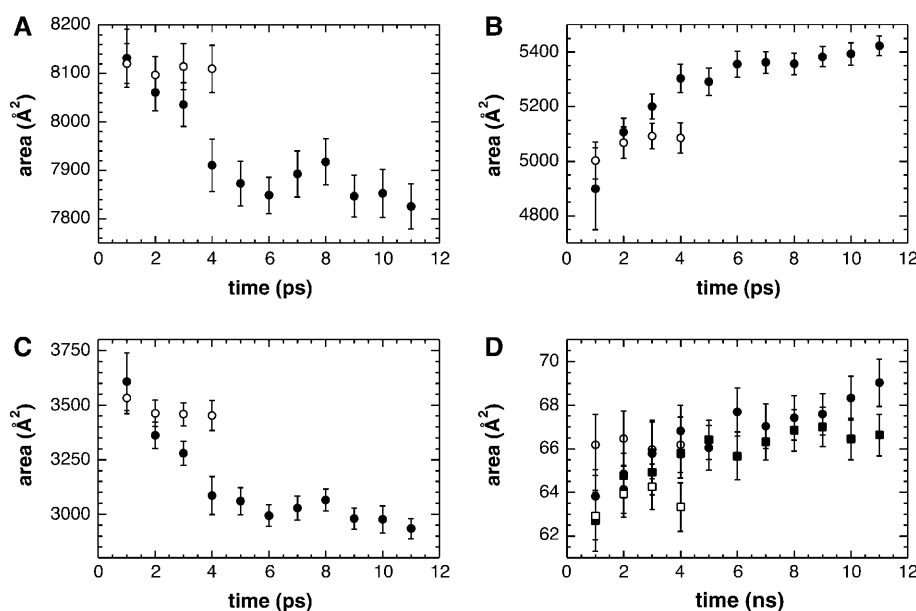


FIGURE 4 Solvent-accessible surface area along the B- to A-DNA transition. One-nanosecond averages, mean  $\pm$  SD are shown as a function of time. (A) Surface area for the complete oligomer; (B) minor groove surface area; (C) major groove surface area; and (D) sugar surface area (purines in circles and pyrimidines in squares). Solid symbols, native simulation; open symbols, control simulation.



proposed that A-DNA is not stable at high humidity because its surface is more hydrophobic than that of B-DNA on account of sugar moieties being more exposed to solvent in A-DNA than in B-DNA (Sprous et al., 1998). Fig. 4D shows that sugar solvent-accessible surface area indeed increases with the transition, both for purines and pyrimidines; most of this area is hydrophobic (only the O4' atom is an H-bond acceptor, and it contributes <10% of the total area). The control simulation does not display a decrease in overall solvent-accessible area, the minor groove does not increase in area, and the major groove compresses minimally. As expected from the distribution of pseudo-rotation angles for purines and pyrimidines, the increase in sugar hydrophobic area is modest for this simulation.

In summary, the simulation appears to capture the essential energetics, and the experimentally and theoretically suggested geometric changes of the B- to A-DNA transition, despite being driven by an as-yet unidentified feature of the CHARMM22 force field. Most importantly, the *slide-first, roll-later* mechanism proposed from a series of independently determined DNA crystal structures appears to be dynamically sensible (see Fig. 3), so this simulation protocol could be useful to derive structures for intermediates along this transition. This suggests that global trends in hydration patterns derived from this simulation are plausible, and thus worth analyzing in detail.

### The evolution of hydration patterns on the surface of DNA

The number of H-bonds formed between all polar DNA atoms and water was calculated as a function of time for the B- to A-DNA transition in 1-ns intervals, and also for the fixed DNA simulations. The number of H-bonds found in the fixed B-DNA simulation was taken as a reference value. The number of H-bonds found in each 1-ns interval, as well as the number found in the fixed A-DNA simulation, was normalized to the value found in the fixed B-DNA simulation. Fig. 5 shows that the number of H-bonds did not change with time or DNA conformation, thus explaining why the DNA-water interaction energy is also invariant along the transition (see above). To see if this was due to compensation between different atom types, the number of H-bonds for each atom type was plotted against time (data not shown), and revealed that this number remains unchanged for all tested atom types. This further allowed for averaging over the whole simulation to produce estimates of hydration numbers for each atom type. These are gathered in Table 1, and agree with the crystallographic analysis performed by Berman and Schneider (1999). The number of H-bonds per basepair is also invariant along the transition, and results in  $19 \pm 1$  H-bonds per basepair, in agreement with volumetric estimates (Chalikian et al., 1999). The simulation has produced first-order hydration properties in agreement with experimental data.

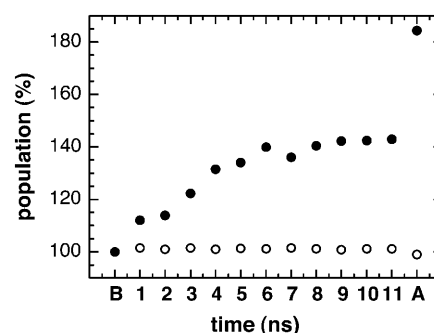


FIGURE 5 H-bonds and water bridges during the B- to A-DNA transition. (A) Corresponds to the fixed A-DNA simulation. Normalized with respect to the fixed B-DNA simulation (B). Average H-bonds and water bridges found in 1-ns intervals. (H-bonds in open circles and water bridges in solid circles.)

Also shown in Fig. 5 is the number of water bridges formed on the surface of DNA. This number is also normalized to that found in the fixed B-DNA simulation. It is clear that although the number of H-bonds did not change, water reorganized on the surface of DNA to form more water bridges. This finding is in agreement with the principle of economy of hydration. The large difference between the last nanosecond and the fixed A-DNA simulation is due to the formation of cross-strand phosphate-phosphate bridges in the latter; these bridges are inefficiently formed in freely moving DNA. Also note that the number of water bridges seems to take  $\sim 6$  ns to equilibrate, compared to the 4 ns required to stabilize DNA geometry. The delay in stabilization of hydration patterns compared to the stabilization of DNA geometry has been noted in other simulations (Cheatham and Kollman, 1997b; Reddy et al., 2003; Trantirek et al., 2000). The increase in water bridges at the surface of DNA is made at the expense of water-water H-bonds. These were calculated for three 1-ns intervals: 1, 6, and 11. In absolute terms, the total number of water bridges formed on DNA increased by 27 from the first to the last nanoseconds, corresponding to an  $\sim 40\%$  increase (see

TABLE 1 Number of H-bonds for each polar atom type

Atom	Number of H-bonds
O1P	$2.60 \pm 0.02$
O2P	$2.73 \pm 0.01$
O5'	$0.47 \pm 0.02$
O3'	$0.71 \pm 0.04$
O4'	$0.71 \pm 0.01$
O2	$0.75 \pm 0.01$
N3	$0.79 \pm 0.04$
H22	$0.69 \pm 0.02$
O4	$0.78 \pm 0.03$
H41	$0.84 \pm 0.01$
O6	$0.88 \pm 0.09$
H61	$0.69 \pm 0.03$
N7	$0.94 \pm 0.02$

Fig. 5); in the same time interval, the number of water-water H-bonds decreased by 31. On average, each bulk water molecule makes 3.4 H-bonds to other water molecules.

An unbiased search for water bridges was carried out; the bridges were originally classified according to the atom types involved in the bridge, resulting in 42 different kinds of bridges representing all of the combinations suggested by Young et al. (1997) for ion coordination in the grooves of DNA. Most of the bridges appear at very low frequencies during the simulation. To distinguish whether the frequency was low because the bridges could be formed at very few locations, or because they were unfavorable due to geometric limitations, these bridges were further classified according to the direction of bridge formation, and whether they occurred along one strand or crossed both strands of DNA. Examination of the DNA oligomer established the number of places where each type of bridge could occur, and the observed frequency from the simulation was normalized by the number of places compatible with that particular kind of bridge. Any bridge with a normalized frequency  $>10\%$  was deemed interesting; 16 such bridges were identified. The time evolution of the frequencies of each of these bridges was calculated, and this resulted in six bridges whose populations (or frequencies) changed concomitantly with the B- to A-DNA transition. These are shown in Fig. 6.

Fig. 6 A shows the increase in water bridges at the DNA backbone, involving the atom types proposed by Saenger et al. (1986), namely, O1P and O2P, and another bridge not discussed by him that involves O1P and O5' of the preceding nucleotide. A snapshot showing these three bridges is shown in Fig. 7. The populations are normalized with respect to the number of sites capable of forming such bridges (the actual number of water bridges is small; in the best case, only one in three possible sites is occupied by a water bridge).

Fig. 6 B depicts the change in water-bridge populations at the minor groove. In agreement with crystallographic analyses, there is a decrease in the number of bridges linking the minor groove edge of the bases to the sugar O4' atom, but they do not disappear completely at the end of the transition. A surprising finding is that N3–O2 cross-strand bridges increase in number as the transition proceeds, with comparable efficiency to O1P–O2P bridges (see Fig. 6 A). These bridges have been associated with the spine of hydration present in A-tracts, intimately linked to narrow minor grooves. The simulated oligomer presents an A-tract, and other opportunities as well, for the formation of this kind of bridge. The result is not an artifact of the simulation: N3–O2 distances in A- and B-DNA are indistinguishable from each other, so there is no reason to assume that these bridges cannot be formed. Possible reasons for these bridges not having been described from crystallographic studies of DNA hydration include the fact that AA/TT steps do not crystallize as A-DNA, and that the wide and exposed minor groove is a common site for duplex-duplex contacts, resulting in incomplete hydration layers in the minor groove.

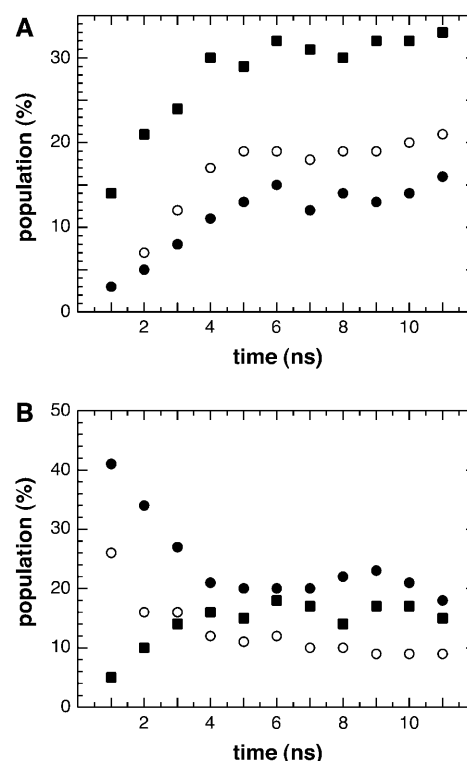


FIGURE 6 Water-bridge populations that change during the B- to A-DNA transition. Normalized to the total number of possible sites forming each particular type of water bridge. (A) Backbone bridges (O1P–O1P in solid circles, O1P–O2P in open circles, and O1P–O5' in solid squares). (B) Minor groove bridges (N3–O4' in solid circles, O2–O4' in open circles, and N3–O2 in solid squares).

One trivial reason for the low abundance of water bridges is that the atoms involved in the bridge are not solvent-accessible. This hypothesis was tested for all polar DNA atoms, by calculating their solvent-accessible surface area as a function of time. Most of these atoms show an increase in solvent accessibility as the transition progresses (for example, O3' as described previously by Lu et al., 2000), or maintain the same area (for example, O2P). Notable exceptions are O1P and O5', whose solvent accessibility decreases during the first 4 ns. As shown above, these atom types are involved in the most common water bridge on the surface of DNA, indicating that as little as  $1 \text{ \AA}^2$  of exposed surface is enough to form water bridges. Another reason for the low population of water bridges is competition from sodium atoms. To test this, the number of sodium ions within  $3 \text{ \AA}$  of polar atoms was counted as a function of time. Only base carbonyls (O2, O4, and O6) contact sodium in a first-shell geometry; phosphate groups interact with hydrated sodium (data not shown). Therefore, competition from sodium cannot explain the dearth of water bridges at the DNA backbone, and it has local effects only over particular bridges in both major and minor grooves, which do not show a correlation to the B- to A-DNA transition.

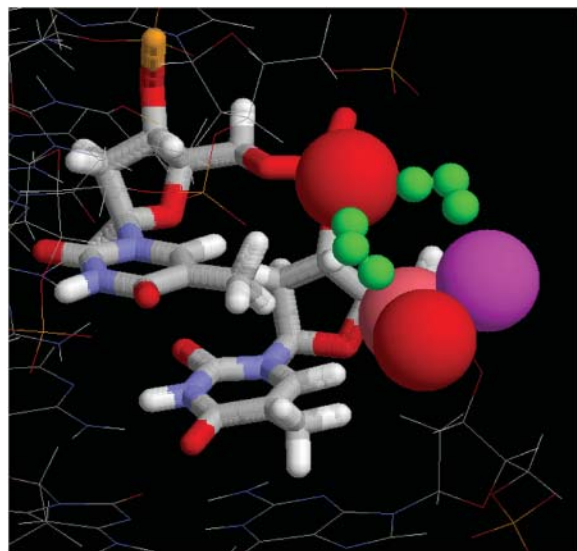


FIGURE 7 Water bridges formed by O1P. Snapshot containing the three bridges in Fig. 6 A; all atoms in standard CPK coloring except for O5' in pink and O2P in purple. Water molecules involved in the bridges colored green. One water molecule is making a three-point contact involving O1P from nucleotide *i*, O1P from nucleotide *i-1*, and O5' from nucleotide *i-1*. The other water-molecule bridges O1P from nucleotide *i* to O2P from nucleotide *i-1*.

Cheatham and Kollman (1997b) suggested that hydration lifetimes depend on the rigidity of the substrate, and that A-type helices are more rigid than B-type helices. Their simulations suggest that the more flexible structures have fewer associated water molecules and counterions. The A-form structures seem to have more specific hydration in the grooves, and so they suggest that this could also participate in rigidity. Fig. 8 A shows the time-dependence of the mean atomic fluctuations of DNA polar atoms. In agreement with the reports by Cheatham and Kollman (1997b) and Lu et al. (2000), average positional fluctuations are  $\sim 20$ –30% higher in B-DNA than in A-DNA. In this case, though, fluctuations converge faster than hydration, suggesting that stable hydration patterns are a consequence of rigidity, and not the other way around. Water bridges will form depending on the availability of ligands at the appropriate distances. Large fluctuations will result in smaller probability of achieving such distances, and therefore, in a smaller population of water bridges.

The role of atomic fluctuations in the stability of hydration patterns is particularly clear for the phosphate-phosphate bridges. Fig. 8 B shows the correlation between O1P fluctuations and the average time spent by a water molecule in one of the three kinds of bridges in which this atom can engage: the fewer the fluctuations (that is, the closer to A-DNA), the larger the average time. In the fixed A-DNA simulation, the average time for the O1P–O1P bridges rises to  $\sim 240$  ps. For the flexible DNA simulation, it is interesting to note that the average time spent by water molecules in making both

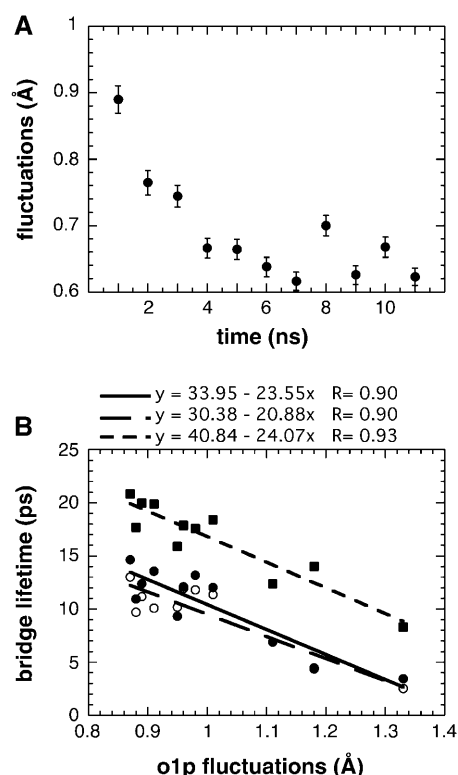


FIGURE 8 Atomic fluctuations as a function of time, and their relation to water-bridge lifetime. (A) Average atomic fluctuations of all polar atoms, in 1-ns intervals mean  $\pm$  SE, as a function of time. (B) Correlation between O1P fluctuations and the average lifetime of the three water bridges shown in Fig. 6 A. (O1P–O1P in solid circles and solid line; O1P–O2P in open circles and dashed line; and O1P–O5' in solid squares and dashed line.)

H-bonds and water bridges is consistent with recent fluorescence spectroscopy measurements of lifetimes of H-bonds (Pal et al., 2003a,b), which suggest that the tightest bound water molecules spend  $\sim 20$  ps on the surface of DNA.

Table 2 summarizes the maximum lifetimes found for the 16 types of water bridges with populations above 10% throughout the simulation. The specific values are comparable to maximum lifetimes reported in the literature (Auffinger and Westhof, 2000, 2001; Bonvin et al., 1998), in simulations done with other force fields (AMBER95 and GROMACS).

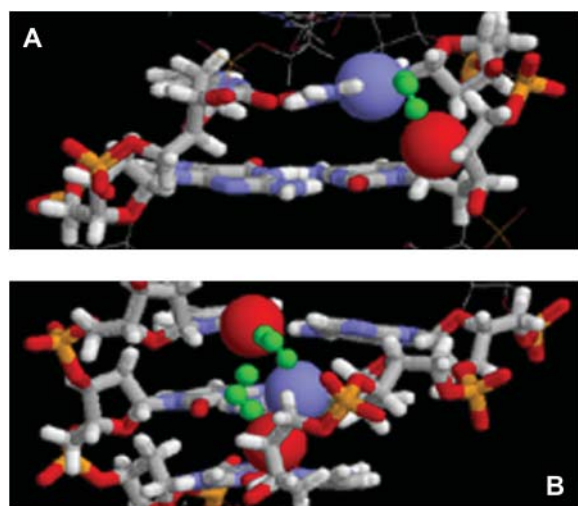
The mechanism responsible for the low abundance of water bridges in the minor groove seems to be different, as the atoms involved in these bridges have smaller atomic fluctuations than the backbone atoms (N3,  $0.43 \pm 0.04$ ; O2,  $0.50 \pm 0.04$ ; and O4',  $0.56 \pm 0.05$ , compared to the *x* axis of Fig. 8 B). In this case, competition between equivalent binding sites seems to be the reason. Although it is possible to form O1P–O1P, O1P–O2P, and O1P–O5' bridges simultaneously (see Fig. 7), it is not possible for an N3 atom to simultaneously engage in N3–O2 and N3–O4' bridges. Fig. 9 shows two snapshots, depicting the same water molecule in two types of interactions with DNA: In Fig. 9 A, it is making a bridge between N3 and O4' in the same strand, whereas in Fig. 9 B N3 is making a water bridge to O2 with another water molecule, and the

**TABLE 2** Maximum lifetime for diverse water bridges/water

	Bridge	Instantaneous lifetime (ps)	Total lifetime (ps)
Backbone	O1P–O1P	80.2	628.1
	O1P–O2P	42.1	324.7
	O1P–O5'	23.8	528.2
Minor groove			
Single-strand	N3–O4'	20.5	439.9
	O2–O4'	19.5	396.8
Cross-strand	N3–O2	10.2	161.7
	O2–O2	11.2	107.6
Major groove			
Intrabase	N7–H61	8.3	473.0
	N7–O6	10.9	643.9
Single-strand	N7–H41	8.5	299.5
	N7–H61	11.8	626.7
	N7–N7	24.5	1096.0
	O4–O6	9.7	100.5
	O6–O6	6.5	211.4
Cross-strand	O4–O6	17.6	213.9
	O6–O6	9.1	218.4

The total lifetime was calculated assuming that each water molecule is in the vicinity of DNA only once during the simulation, and disregarding the time interval between rebinding events. The instantaneous lifetime is calculated from the longest uninterrupted time span that a water molecule forms a water bridge to the listed atoms.

water molecule bound to O4' forms an H-bond to the N3–O2 bridging water. In this version of the CHARMM potential, the partial charges are  $0.4e$  for O4', and  $0.47e$  for O2, indicating that these two sites are practically equivalent energetically. This equivalence may lead to frustration, and consequently, to short bridge lifetimes.



**FIGURE 9** Competition between equivalent sites at the minor groove. All atoms in standard CPK colors, except water, which is shown in green. N3, O2, and O4' atoms involved in the bridges are shown in space-filling mode. (A) N3–O4' bridge. (B) N3–O2 bridge, and water-water bridge between a water molecule H-bonded to O4' and the one making the N3–O2 bridge.

## SUMMARY AND CONCLUSIONS

The B- to A-DNA transition presented here, even though driven by a force-field bias, appears to be mechanistically reasonable: it seems to be driven in this version of the force field by sodium condensation, *slide* stabilizes before *roll*, there is a nearly perfect correlation between  $\delta$  and  $\chi$ , and G-tracts undergo the transition first. The structural features of the intermediate structures described crystallographically are reproduced by the simulation, indicating that the intermediate structures are mechanistically and dynamically feasible. Therefore, the simulation protocol presented here could be useful to generate plausible structures for intermediates thought to be the recognition sites of DNA binding proteins (Lu et al., 2000).

Water bridges represent the most economical way to hydrate the surface of molecules, with an associated entropic cost. They further can be thought of as molecular rulers, since they will only be present when the distance between the bridged atoms is adequate, thus becoming sensitive monitors of both the structure and the dynamics of the macromolecule to which they are bound. The analysis presented here would suggest that hydration patterns are a consequence of DNA geometry and motion, instead of the driving force for the conformational transition. The analysis of the evolution of hydration patterns along the transition allowed for the identification of water bridges associated with the principle of economy of hydration, and of an unexpected bridge in the minor groove. The presence or absence of water bridges on the surface of DNA depends on the interplay between achieving the correct distance among the acceptor atoms, the local flexibility, and the competition with nearby acceptors. This interplay may be related to the experimentally determined short lifetimes of water-DNA H-bonds, and to the role of water as a lubricant in DNA-ligand interactions.

## SUPPLEMENTARY MATERIAL

An online supplement to this article can be found by visiting BJ Online at <http://www.biophysj.org>. This pertains to the time evolution of the phase angle for purines and pyrimidines for both the native and control simulations.

The author thanks Harel Weinstein for critical reading of the manuscript, and helpful suggestions.

The author thanks the program “C  mputo Cient  fico” (SEP-FOMES 2000) for unlimited computer access to the IBM-4 Regatta at Universidad Aut  noma del Estado de Morelos, and also thanks the Mount Sinai School of Medicine for unlimited computer access. Funded by Consejo Nacional de Ciencia y Tecnolog  a (No. J33190-E).

## REFERENCES

- Auffinger, P., and E. Westhof. 2000. Water and ion binding around RNA and DNA (C,G) oligomers. *J. Mol. Biol.* 300:1113–1131.
- Auffinger, P., and E. Westhof. 2001. Water and ion binding around r(UpA)<sup>12</sup> and d(TpA)<sup>12</sup> oligomers—comparison with RNA and DNA (CpG)<sup>12</sup> duplexes. *J. Mol. Biol.* 305:1057–1072.

- Basham, B., G. P. Schroth, and P. S. Ho. 1995. An A-DNA triplet code: thermodynamic rules for predicting A- and B-DNA. *Proc. Natl. Acad. Sci. USA*. 92:6464–6468.
- Berman, H. M., and B. Schneider. 1999. Nucleic acid hydration. In *Oxford Handbook of Nucleic Acid Structure*, 1st Ed. Oxford University Press, New York. 295–312.
- Bonvin, A. M. J. J., M. Sunnerhagen, G. Otting, and W. F. van Gunsteren. 1998. Water molecules in DNA recognition. II. A molecular dynamics view of the structure and hydration of the Trp operator. *J. Mol. Biol.* 282:859–873.
- Brooks, B. R., R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. 1983. CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* 4:187–217.
- Chalikian, T. V., J. Volker, A. R. Srinivasan, W. K. Olson, and K. J. Breslauer. 1999. The hydration of nucleic acid duplexes as assessed by a combination of volumetric and structural techniques. *Biopolymers*. 50:459–471.
- Cheatham III, T. E., and P. A. Kollman. 1997a. Insight into the stabilization of A-DNA by specific ion association: spontaneous B-DNA to A-DNA transitions observed in molecular dynamics simulations of d[ACCCGCGGGT]<sup>2</sup> in the presence of hexaamminecobalt(III). *Structure*. 5:1297–1311.
- Cheatham III, T. E., and P. A. Kollman. 1997b. Molecular dynamics simulations highlight the structural differences among DNA:DNA, RNA:RNA, and DNA:RNA hybrid duplexes. *J. Am. Chem. Soc.* 119:4805–4825.
- Cheatham III, T. E., J. Srinivasan, D. A. Case, and P. A. Kollman. 1998. Molecular dynamics and continuum solvent studies of the stability of polyG-polyC and polyA-polyT DNA duplexes in solution. *J. Biomol. Struct. Dyn.* 16:265–280.
- Cornell, W. D., P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, Jr., D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman. 1995. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* 117:5179–5197.
- Dickerson, R. E., and H. L. Ng. 2001. DNA structure from A to B. *Proc. Natl. Acad. Sci. USA*. 98:6986–6988.
- Dornberger, U., N. Spackovj, A. Walter, F. A. Gollmick, J. Sponer, and H. Fritzsche. 2001. Solution structure of the dodecamer d(CATGGGCC-CATG)<sup>2</sup> is B-DNA. Experimental and molecular dynamics study. *J. Biomol. Struct. Dyn.* 19:159–174.
- Feig, M., and B. M. Pettitt. 1997. Experiment vs. force fields: DNA conformation from molecular dynamics simulations. *J. Phys. Chem. B*. 101:7361–7363.
- Feig, M., and B. M. Pettitt. 1998. Structural equilibrium of DNA represented with different force fields. *Biophys. J.* 75:134–149.
- Feller, S. E., Y. Zhang, R. W. Pastor, and B. R. Brooks. 1995. Constant pressure molecular dynamics simulation: the Langevin piston method. *J. Chem. Phys.* 103:4613–4621.
- Foloppe, N., and A. D. MacKerell, Jr. 1999. Intrinsic conformational properties of deoxyribonucleosides: implicated role for cytosine in the equilibrium among the A, B, and Z forms of DNA. *Biophys. J.* 76:3206–3218.
- Franklin, R. E., and R. G. Gosling. 1953. Molecular configuration in sodium thymonucleate. *Nature*. 171:740–741.
- Jayaram, B., D. Sprous, M. A. Young, and D. L. Beveridge. 1998. Free energy analysis of the conformational preferences of A and B forms of DNA in solution. *J. Am. Chem. Soc.* 120:10629–10633.
- Jones, S., P. van Heyningen, H. M. Berman, and J. M. Thornton. 1999. Protein-DNA interactions: a structural analysis. *J. Mol. Biol.* 287:877–896.
- Jorgensen, W. L., J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935.
- Jose, D., and D. Porschke. 2004. Dynamics of the B-A transition of DNA double helices. *Nucleic Acids Res.* 32:2251–2258.
- Lee, B., and F. M. Richards. 1971. The interpretation of protein structures: estimation of static accessibility. *J. Mol. Biol.* 55:379–400.
- Lu, X. J., Z. Shakked, and W. K. Olson. 2000. A-form conformational motifs in ligand-bound DNA structures. *J. Mol. Biol.* 300:819–840.
- MacKerell, A. D., Jr., N. Banavali, and N. Foloppe. 2001. Development and current status of the CHARMM force field for nucleic acids. *Biopolymers*. 56:257–265.
- MacKerell, A. D., Jr., J. Wiorkiewicz-Kuczera, and M. Karplus. 1995. An all-atom empirical energy function for the simulation of nucleic acids. *J. Am. Chem. Soc.* 117:11946–11975.
- Manning, G. S. 1978. The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides. *Q. Rev. Biophys.* 11:179–246.
- Mazur, A. K. 2003. Titration in silico of reversible B ↔ A transitions in DNA. *J. Am. Chem. Soc.* 125:7849–7859.
- Mehrotra, P. K., and D. L. Beveridge. 1980. Structural analysis of molecular solutions based on quasi-component distribution functions. Application to [H<sub>2</sub>CO]<sup>aq</sup> at 25°C. *J. Am. Chem. Soc.* 102:4287–4294.
- Mehrotra, P. K., F. T. Marchese, and D. L. Beveridge. 1981. Statistical state solvation sites. *J. Am. Chem. Soc.* 103:672–673.
- Nekludova, L., and C. O. Pabo. 1994. Distinctive DNA conformation with enlarged major groove is found in Zn-finger-DNA and other protein-DNA complexes. *Proc. Natl. Acad. Sci. USA*. 91:6948–6952.
- Ng, H. L., and R. E. Dickerson. 2002. Mediation of the A/B-DNA helix transition by G-tracts in the crystal structure of duplex CATGGGCC-CATG. *Nucleic Acids Res.* 30:4061–4067.
- Ng, H. L., M. L. Kopka, and R. E. Dickerson. 2000. The structure of a stable intermediate in the A ↔ B DNA helix transition. *Proc. Natl. Acad. Sci. USA*. 97:2035–2039.
- Norberg, J., and L. Nilsson. 2000. On the truncation of long-range electrostatic interactions in DNA. *Biophys. J.* 79:1537–1553.
- Olson, W. K., A. A. Gorin, X. J. Lu, L. M. Hock, and V. B. Zhurkin. 1998. DNA sequence-dependent deformability deduced from protein-DNA crystal complexes. *Proc. Natl. Acad. Sci. USA*. 95:11163–11168.
- Pal, S. K., L. Zhao, T. Xia, and A. H. Zewail. 2003a. Site- and sequence-selective ultrafast hydration of DNA. *Proc. Natl. Acad. Sci. USA*. 100:13746–13751.
- Pal, S. K., L. Zhao, and A. H. Zewail. 2003b. Water at DNA surfaces: ultrafast dynamics in minor groove recognition. *Proc. Natl. Acad. Sci. USA*. 100:8113–8118.
- Parsegian, V. A., R. P. Rand, and D. C. Rau. 1995. Macromolecules and water: probing with osmotic stress. *Methods Enzymol.* 259:43–94.
- Ravishanker, G., S. Swaminathan, D. L. Beveridge, R. Lavery, and H. Sklenar. 1989. Conformational and helicoidal analysis of 30 PS of molecular dynamics on the d(CGCGAATTCGCG) double helix: “curves,” dials, and windows. *J. Biomol. Struct. Dyn.* 6:669–699.
- Reddy, S., F. Leclerc, and M. Karplus. 2003. DNA polymorphism: a comparison of force fields for nucleic acids. *Biophys. J.* 84:1421–1449.
- Ryckaert, J. P., G. Ciccotti, and H. J. C. Berendsen. 1977. Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes. *J. Comput. Phys.* 23:327–341.
- Saenger, W., W. N. Hunter, and O. Kennard. 1986. DNA conformation is determined by economics in the hydration of phosphate groups. *Nature*. 324:385–388.
- Sprous, D., M. A. Young, and D. L. Beveridge. 1998. Molecular dynamics studies of the conformational preferences of a DNA double helix in water and an ethanol/water mixture: theoretical considerations of the A ↔ B transition. *J. Phys. Chem. B*. 102:4658–4667.
- Srinivasan, J., T. E. Cheatham III, P. Cieplak, P. A. Kollman, and D. A. Case. 1998. Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate-DNA helices. *J. Am. Chem. Soc.* 120:9401–9409.

- Tolstorukov, M. Y., V. I. Ivanov, G. G. Malenkov, R. L. Jernigan, and V. B. Zhurkin. 2001. Sequence-dependent B $\leftrightarrow$ A transition in DNA evaluated with dimeric and trimeric scales. *Biophys. J.* 81:3409–3421.
- Tonelli, M., E. Ragg, A. M. Bianucci, K. Lesiak, and T. L. James. 1998. Nuclear magnetic resonance structure of d(GCATATGATAG). d(CTATCATATGC): a consensus sequence for promoters recognized by  $\sigma$ K RNA polymerase. *Biochemistry*. 37:11745–11761.
- Trantirek, L., R. Stefl, M. Vorlickova, J. Koca, V. Sklenar, and J. Kypr. 2000. An A-type double-helix of DNA having B-type puckering of the deoxyribose rings. *J. Mol. Biol.* 297:907–922.
- Vargason, J. M., K. Henderson, and P. S. Ho. 2001. A crystallographic map of the transition from B-DNA to A-DNA. *Proc. Natl. Acad. Sci. USA*. 98:7265–7270.
- Winger, R. H., K. R. Liedl, A. Pichler, A. Hallbrucker, and E. Mayer. 2000. B-DNA's BII conformer substate population increases with decreasing water activity. 1. A molecular dynamics study of d(CGCGAATTCGCG)<sup>2</sup>. *J. Phys. Chem. B*. 104:11349–11353.
- Winger, R. H., K. R. Liedl, S. Rüdiger, A. Pichler, A. Hallbrucker, and E. Mayer. 1998. B-DNA's BI  $\rightarrow$  BII conformer substrate dynamics in coupled with water migration. *J. Phys. Chem. B*. 102:8934–8940.
- Yang, L., and B. M. Pettitt. 1996. B to A transition of DNA on the nanosecond time scale. *J. Phys. Chem.* 100:2564–2566.
- Young, M. A., B. Jayaram, and D. L. Beveridge. 1997. Intrusion of counterions into the spine of hydration in the minor groove of B-DNA: fractional occupancy of electronegative pockets. *J. Am. Chem. Soc.* 117: 59–69.